

Concepts of a quantum information theory of many letters

Kim J. Boström

Institut für Physik, Universität Potsdam, 14469 Potsdam, Germany
(Split version 2.1 beta / February 1, 2008)

A theoretical framework is presented allowing the treatment of quantum messages with components of variable length. To this aim a many-letter space, similar to the Fock space, is constructed, generalizing the standard quantum information theory of block messages of fixed length. In the many-letter space a length operator can be defined measuring the length of a quantum message, whose eigenspaces are the block Hilbert spaces used in the standard theory.

I. INTRODUCTION

Information theory is the theory of messages composed from letters. In classical information theory a message is represented by the state of a classical system composed of many subsystems representing the letters of the message. Quantum information theory is much the same, though here the systems are *quantum*. Since quantum systems obey the laws of quantum mechanics, the situation is radically different from the classical case. However, whereas in classical information theory there is no difficulty in dealing with messages of variable length, quantum information theory, which is usually based on Hilbert spaces of a fixed dimension, does not allow a simple treatment of quantum messages in a superposition of distinct lengths. In this paper a theoretical framework is presented which allows the treatment of such quantum messages in an intuitive way. It is based on the notion of an infinite *direct sum* of Hilbert spaces, which obtains physical meaning if one imagines e.g. a source of photons whose number is a quantum mechanical observable, i.e. the state of the photon ray is generally in a *superposition* of distinct photon number states. The only difference between particles and quantum letters is that the letter systems are *distinguishable*. So the many-letter quantum information theory presented here is just many-particle quantum mechanics with distinguishable particles. Within this framework a close analogy between concepts of classical and quantum information theory can be established, while the standard quantum information theory is fully contained.

This paper is separated into two parts. The first part reviews roughly some basic concepts of classical information theory in order to motivate the corresponding notions presented in the second part, which is dedicated to quantum information theory. A detailed summary of classical information theory can be found in [1], a very recommendable review on quantum information theory is given in [2].

II. CLASSICAL MESSAGES

A. General messages and block messages

The basic object in information theory is a message. A *classical message* is a string \mathbf{x} of letters x taken from an alphabet \mathcal{A} of size $|\mathcal{A}|$ and is denoted by $\mathbf{x} = (x_1 \cdots x_n)$. Let us denote strings of length n explicitly by

$$x^n := (x_1 \cdots x_n) \quad . \quad (1)$$

The set of *block messages* x^N of fixed length N is written as

$$\mathcal{A}^N := \{(x_1 \cdots x_N) \mid x_n \in \mathcal{A}\} \quad . \quad (2)$$

Let us also allow for the *empty message* $x^0 = (\cdot)$ that forms the set $\mathcal{A}^0 := \{(\cdot)\}$. The set of all messages of finite length is defined by

$$\mathcal{A}^+ := \bigcup_{n=0}^{\infty} \mathcal{A}^n \quad . \quad (3)$$

Now Alice wants to communicate general messages to Bob. There are certain messages she wants to send and others (perhaps nonsense or too nasty messages) she does not. So she extracts a *source set* $\Omega \subset \mathcal{A}^+$ and sends each message $x^n \in \Omega$ with a *a priori* probability $p(x^n) > 0$. To Bob, who does not know what Alice is doing, the message appears as a random variable \mathbf{X} , defined by the source set Ω and the *a priori* probabilities $p(x^n)$:

$$\mathbf{X} := \{[x^n, p(x^n)] \mid x^n \in \Omega\} \quad , \quad (4)$$

where $p(x^n) > 0$ for all $x^n \in \Omega$ and $\sum_{x^n \in \Omega} p(x^n) = 1$. The random variable \mathbf{X} is called a *message ensemble*.

If Bob wants to analyze the messages, he performs a measurement on a received message x^n and obtains a real number $A(x^n)$. The ensemble average of his *observable* $A : \Omega \rightarrow \mathbb{R}$ is then given by

$$\bar{A}(\mathbf{X}) \equiv \langle A(\mathbf{X}) \rangle := \sum_{x^n \in \Omega} p(x^n) A(x^n) \quad . \quad (5)$$

He may, for example, measure the length of a message, given by the *length function* $L : \mathcal{A}^+ \rightarrow \mathbb{N}$ with $L(x^n) = n$ and the length of the empty message being

set to $L(\cdot) := 0$. The expected length of a message from Alice is then

$$\bar{L}(\mathbf{X}) = \sum_{x^n \in \Omega} p(x^n) n \quad . \quad (6)$$

If \mathbf{X} is a block message, its length is fixed to some N that is known to both Alice and Bob.

B. Canonical messages

There is a type of message that is of fundamental importance to information theory and that is why it is named here the *canonical message*. It is a block message of fixed length N formed by independent identically distributed letters. Alice takes the *letter ensemble*

$$\mathbf{X} := \{[x, p(x)] \mid x \in \mathcal{A}\} \quad (7)$$

and composes messages by just putting N letters in a row, i.e. $\mathbf{X} = X^N := (X_1 \cdots X_N)$, $X_n = X$, resulting in a *canonical message ensemble*

$$X^N = \{[x^N, p(x^N)] \mid x^N \in \mathcal{A}^N\} \quad (8)$$

with $p(x^N) = p(x_1) \cdots p(x_N)$.

III. QUANTUM MESSAGES

Quantum information theory can be obtained straightforwardly by mapping classical objects to quantum objects. To put it simple: *Classical information* is carried by *classical states* of a medium and *quantum information* is carried by *quantum states* of a medium. Imagine Alice writing her messages not on a sheet of paper or imprinting it onto the surface of a magnetic tape or a hard disk, but instead modifying single atoms, molecules, electrons, photons or any other microscopic systems that can only be described by the laws of quantum mechanics. The mathematical framework of classical information theory then has to be translated into the language of quantum mechanics. The result of this quantization procedure is *quantum information theory*.

Alice prepares the medium to be in a quantum state $|\varphi\rangle$, performs some operations on it and sends it to Bob. The message has been successfully transmitted, if Bob, after performing some operations on the received state, ends up with the same state $|\varphi\rangle$ that Alice originally prepared. Note that it is not necessary for Bob to perform any *measurement* on the state. Bob does not need to *know* which state Alice has originally prepared. This would transform the quantum information contained in that state to classical information. It is a major difference between classical and quantum information that *knowledge*, i.e. the state of someone's brain, is always classical, whereas the state of an unknown quantum state is *intrinsically unknowable*, since there is no operation in the world allowing to guess an unknown state with perfect fidelity.

A. Quantum alphabet

1. A priori alphabet

A classical letter is represented by the state of a classical system. If the system is quantum instead, the letter corresponds to a quantum state. Thus a classical letter x can be transformed into a quantum letter by mapping it to a normalized Hilbert vector $|x\rangle \in \mathcal{H}$. In such a way, the classical alphabet \mathcal{A} is mapped to a *quantum alphabet*

$$\mathcal{Q} := \{|x\rangle \in \mathcal{H} \mid x \in \mathcal{A}\} \quad . \quad (9)$$

The Hilbert space spanned by the letters of the quantum alphabet is the *letter space*

$$\mathcal{H}_{\mathcal{Q}} := \text{Span}(\mathcal{Q}) \quad , \quad (10)$$

where its dimension given by $K_{\mathcal{Q}} := \dim \mathcal{H}_{\mathcal{Q}} \leq |\mathcal{Q}| = |\mathcal{A}|$, with equality if the letter states are linearly independent.

The quantum letters in \mathcal{Q} are not required to be mutual orthogonal, yet not even linearly independent. So it is in general not possible for Bob to perfectly *distinguish* the letters that Alice choses from her alphabet \mathcal{Q} , which is thus called an *a priori alphabet*. In sad words: Bob will generally not be able to *read* a message from Alice. Instead he probably recognizes *a posteriori* letters, that also lie in the letter space spanned by the *a priori* alphabet but which are different from the letters that Alice originally had sent. This is a typically quantum phenomenon with no classical correspondance.

2. Basis alphabet

A set $\mathcal{B}_{\mathcal{Q}} = \{|a\rangle\}_a$ of mutually orthogonal normalized basis vectors of the letter space $\mathcal{H}_{\mathcal{Q}}$ is called a *basis alphabet* corresponding to \mathcal{Q} , so

$$\sum_{a \in \mathcal{B}_{\mathcal{Q}}} |a\rangle \langle a| = \mathbb{1}_{\mathcal{H}_{\mathcal{Q}}} \quad \langle a|a'\rangle = \delta_{aa'} \quad . \quad (11)$$

Since all *basis letters* $|a\rangle \in \mathcal{B}_{\mathcal{Q}}$ are perfectly distinguishable, they can be viewed as almost classical. The basis alphabet is a very important concept in quantum information theory. Any single-letter message from Alice can be expressed as a superposition of basis letters:

$$|x\rangle = \sum_{a \in \mathcal{B}_{\mathcal{Q}}} \langle a|x\rangle |a\rangle \quad . \quad (12)$$

If Bob happens to measure along the basis letter subspaces, the *a priori* message from Alice will decohere into its basis letter components. Up to the measurement, though, they are all simultaneously engaged. The number of basis letters equals the dimension of the letter space, so there are probably less basis letters than *a priori* letters, i.e. $|\mathcal{B}_{\mathcal{Q}}| = \dim \mathcal{H}_{\mathcal{Q}} \leq |\mathcal{Q}|$.

B. Block messages

1. General block messages

A classical string $x^n \in \mathcal{A}^n$ of length n , given by $x^n = (x_1 \cdots x_n)$, is mapped to a Hilbert vector $|x^n\rangle \in \mathcal{H}^n$ normalized to unity and formed by the tensor product of the letter states corresponding to the letters contained in x^n . It is called a *product message* or *quantum string* and is denoted by

$$|x^n\rangle := |x_1 \cdots x_n\rangle \equiv |x_1\rangle \otimes \cdots \otimes |x_n\rangle \quad (13)$$

The set \mathcal{A}^N of classical strings of fixed length N is mapped to the set of *quantum block strings*

$$\mathcal{Q}^N := \{|x^N\rangle \in \mathcal{H}^N \mid x^N \in \mathcal{A}^N\} \quad (14)$$

Let us allow also for the empty quantum message $|x^0\rangle = |\cdot\rangle$ that forms the set $\mathcal{Q}^0 := \{|\cdot\rangle\}$. The Hilbert space spanned by the elements of \mathcal{Q}^N is the N -fold tensor product of the letter space and is called the *block message space*:

$$\mathcal{H}_{\mathcal{Q}}^N := \text{Span}(\mathcal{Q}^N) = \mathcal{H}_{\mathcal{Q}} \otimes \cdots \otimes \mathcal{H}_{\mathcal{Q}} \quad , \quad (15)$$

where its dimension is given by $\dim \mathcal{H}_{\mathcal{Q}}^N = (\dim \mathcal{H}_{\mathcal{Q}})^N \leq |\mathcal{Q}|^N$. The one-dimensional empty message space is defined by $\mathcal{H}_{\mathcal{Q}}^0 := \text{Span}(\mathcal{Q}^0)$.

What messages can Alice compose now? She can prepare the quantum string $|x^N\rangle$ of length N by manipulating each of the N letter systems separately. But quantum mechanics allows her also to perform unitary operations on the entire message state $|x^N\rangle \in \mathcal{Q}^N$ before sending it to Bob. So she can construct any normalized vector $|\varphi(x^N)\rangle \in \mathcal{H}_{\mathcal{Q}}^N$ by performing $|\varphi(x^N)\rangle = U(x^N)|x^N\rangle$, where $U(x^N)$ is a unitary operator on $\mathcal{H}_{\mathcal{Q}}^N$. Though $|x^N\rangle$ is a *product message*, $|\varphi(x^N)\rangle$ generally is not. In that case it is an *entangled message*. While quantum strings are always product messages, general block messages can be arbitrary superpositions of strings of the same length. There is no classical correspondence to such objects. In order to make it explicitely we may denote a block message $|\varphi\rangle$ of length N by a small index N like in $|\varphi\rangle_N$. A general form of block messages is then given by

$$|\varphi\rangle_N = \sum_{a^N} \varphi(a^N) |a^N\rangle \quad , \quad (16)$$

where $\varphi(a^N) = \langle a^N | \varphi \rangle$ and $\mathcal{B}_{\mathcal{Q}} = \{a\}_a$ being a set of mutually orthogonal basis letters of the letter space $\mathcal{H}_{\mathcal{Q}}$. The sum is performed over all strings $|a^N\rangle$ over the basis alphabet. By applying her unitary operations to the strings $|x^N\rangle$ of \mathcal{Q}^N , Alice prepares a set of general *block messages* of fixed length N

$$\Gamma := \{|\varphi\rangle_N \in \mathcal{H}_{\mathcal{Q}}^N \mid p(\varphi) > 0\} \quad (17)$$

Alice choses the message $|\varphi\rangle_N \in \Gamma$ with *a priori* probability $p(\varphi)$, i.e. she draws each message from the *message ensemble*

$$|\Phi\rangle_N := \{[|\varphi\rangle_N, p(\varphi)] \mid |\varphi\rangle_N \in \Gamma\} \quad (18)$$

If Bob receives the message $|\varphi\rangle_N$ and tries to get some classical information out of it, he performs a measurement of an observable A , represented by a self-adjoint operator on $\mathcal{H}_{\mathcal{Q}}^N$. Each time he does, he gets a random result whose quantum mechanical expectation value is given by

$$A(\varphi) \equiv \langle A \rangle_{\varphi} := \langle \varphi | A | \varphi \rangle_N \quad (19)$$

Since Alice draws her messages from the ensemble $|\Phi\rangle_N$, the *ensemble average* of A is ruled by

$$A(\Phi) := \langle \Phi | A | \Phi \rangle_N = \sum_{\varphi \in \Gamma} p(\varphi) \langle \varphi | A | \varphi \rangle_N \quad (20)$$

Equivalently, Bob can calculate the ensemble average using the *message matrix*

$$\sigma := \sum_{\varphi \in \Gamma} p(\varphi) |\varphi\rangle \langle \varphi|_N \quad (21)$$

and find the ensemble average being governed by

$$A(\sigma) \equiv \langle A \rangle_{\sigma} := \text{Tr}_N \{ \sigma A \} \quad , \quad (22)$$

where Tr_N denotes the trace over the space $\mathcal{H}_{\mathcal{Q}}^N$. It is a profound peculiarity of quantum mechanics that Bob would end up with the same statistical average, if Alice had taken some other message ensemble yielding the same message matrix σ . Consequently, there is more information in knowing the *ensemble* $|\Phi\rangle$ (like Alice does) than just knowing the *matrix* σ . This additional information is in no way available by performing measurements on the message states. Nevertheless, these two distinct notions are both used within quantum information theory.

To Bob there would be no difference if Alice had taken the message ensemble

$$|E\rangle_N := \{[|e_k\rangle_N, q_k] \mid k = 1 \dots K^N\} \quad , \quad (23)$$

with the $|e_k\rangle_N$'s being the eigenstates of σ ,

$$\sigma = \sum_{k=1}^{K^N} q_k |e_k\rangle \langle e_k|_N \quad , \quad (24)$$

where

$$\langle e_k | e_l \rangle_N = \delta_{kl}, \quad \sum_{k=1}^{K^N} |e_k\rangle \langle e_k|_N = \mathbb{1}_N \quad (25)$$

and $K^N := \dim \mathcal{H}_{\mathcal{Q}}^N = (\dim \mathcal{H}_{\mathcal{Q}})^N$. This special ensemble is a very interesting one, since here the single messages $|e_k\rangle_N$ can be *distinguished* from another by a suitable measurement. So it is the most classical equivalent ensemble corresponding to what Alice is doing.

2. Product block messages

Alice prepares her message letter by letter and obtains a product state $|x^N\rangle = |x_1 \cdots x_N\rangle \in \mathcal{H}_{\mathcal{Q}}^N$. She prepares the state $|x^N\rangle$ with *a priori* probability $p(x^N)$, i.e. she draws her messages from the *product message ensemble*

$$|X^N\rangle := \{|x^N\rangle, p(x^N)\} \mid x^N \in \Omega \subset \mathcal{A}^N \quad (26)$$

Now the corresponding message matrix,

$$\sigma = \sum_{x_1 \cdots x_N} p(x_1 \cdots x_N) \left[|x_1\rangle\langle x_1| \otimes \cdots \otimes |x_N\rangle\langle x_N| \right] \quad (27)$$

$$= \left[\sum_{x_1} p_1(x_1) |x_1\rangle\langle x_1| \right] \otimes \cdots \otimes \left[\sum_{x_N} p_N(x_N) |x_N\rangle\langle x_N| \right] \quad (28)$$

falls apart into a product $\sigma = \rho_1 \otimes \cdots \otimes \rho_N$ of *single-letter matrices* ρ_n , given by

$$\rho_n := \sum_{x_n \in \mathcal{A}} p_n(x_n) |x_n\rangle\langle x_n| \quad (29)$$

with the *marginal probabilities*

$$p_n(x_n) := \sum_{x_i: i \neq n} p(x_1 \cdots x_N) \quad (30)$$

Again it is interesting to regard the spectral decomposition of each single-letter matrix,

$$\rho_n = \sum_{k=1}^K q_{nk} |e_{nk}\rangle\langle e_{nk}| \quad (31)$$

with

$$\langle e_{nk} | e_{nl} \rangle = \delta_{kl}, \quad \sum_{k=1}^K |e_{nk}\rangle\langle e_{nk}| = \mathbb{1}_{\mathcal{H}_{\mathcal{Q}}} \quad (32)$$

and $K := \dim \mathcal{H}_{\mathcal{Q}}$. To Bob it appears as if Alice had prepared messages over orthogonal basis alphabets $\mathcal{B}_n := \{|e_{nk}\rangle, q_{nk}\} \mid k = 1 \dots K$, that vary from letter to letter.

3. Canonical messages

Canonical messages are product block messages $|x^N\rangle \in \mathcal{H}_{\mathcal{Q}}^N$ over an *a priori* alphabet $\mathcal{Q} = \{|x\rangle\}_x$, chosen with factorizing *a priori* probabilities $p(x^N) = p(x_1) \cdots p(x_N)$. The message matrix,

$$\sigma = \rho^{\otimes N} = \rho \otimes \cdots \otimes \rho \quad (33)$$

is the N -fold tensor product of the *letter matrix*,

$$\rho = \sum_x p(x) |x\rangle\langle x| \quad (34)$$

Alice uses an *a priori* letter ensemble

$$|X\rangle = \{|x\rangle, p(x)\} \mid |x\rangle \in \mathcal{Q} \quad (35)$$

which is just put in a row N times to form the *canonical ensemble* $|X^N\rangle$. To Bob there is no difference if Alice instead uses the *basis letter ensemble* $|A\rangle$ consisting of the ρ eigenstates $|a\rangle$, i.e.

$$\rho = \sum_a q(a) |a\rangle\langle a| \quad (36)$$

and forms the message ensemble $|A^N\rangle$.

IV. MANY-LETTER MESSAGES

A. General many-letter messages

Standard quantum information theory describes only block messages. We like to go further now and allow quantum messages of arbitrary length. To this aim we seek a quantum analog of the set \mathcal{A}^+ of classical messages of arbitrary length. It is easily found by mapping each classical message $x^n \in \mathcal{A}^+$ to a product Hilbert vector $|x^n\rangle \in \mathcal{H}^n$. Regard the set of product block messages of length n ,

$$\mathcal{Q}^n := \{|x^n\rangle \in \mathcal{H}^n \mid x^n \in \mathcal{A}^n\} \quad (37)$$

with $\mathcal{Q}^0 := \{|\cdot\rangle\}$ being defined as the set formed by the empty message $|\cdot\rangle$. The Hilbert space spanned by the members of \mathcal{Q}^n is given by

$$\mathcal{H}_{\mathcal{Q}}^n := \text{Span}(\mathcal{Q}^n) \quad (38)$$

Now construct the infinite set

$$\mathcal{Q}_+ := \bigcup_{n=0}^{\infty} \mathcal{Q}^n \quad (39)$$

The space spanned by the elements of \mathcal{Q}_+ (regarding that messages of distinct length are always orthogonal) is the *many-letter space*

$$\mathcal{M}_{\mathcal{Q}} := \bigoplus_{n=0}^{\infty} \mathcal{H}_{\mathcal{Q}}^n \quad (40)$$

The *direct sum* of two Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2$ is defined as the orthogonal sum of their elements, i.e.

$$\begin{aligned} \mathcal{H}_1 \oplus \mathcal{H}_2 \\ := \{|\psi_1\rangle_1 + |\psi_2\rangle_2 \mid |\psi_1\rangle_1 \in \mathcal{H}_1, |\psi_2\rangle_2 \in \mathcal{H}_2\}, \end{aligned} \quad (41)$$

and is a Hilbert space with the scalar product

$$\begin{aligned} & \left({}_1\langle\psi_1| + {}_2\langle\psi_2| \right) \left(|\varphi_1\rangle_1 + |\varphi_2\rangle_2 \right) \\ & \quad := \langle\psi_1|\varphi_1\rangle_1 + \langle\psi_2|\varphi_2\rangle_2 \quad , \end{aligned} \quad (42)$$

and the dimension $\dim(\mathcal{H}_1 \oplus \mathcal{H}_2) = \dim \mathcal{H}_1 + \dim \mathcal{H}_2$. Both spaces \mathcal{H}_1 and \mathcal{H}_2 are orthogonal subspaces of $\mathcal{H}_1 \oplus \mathcal{H}_2$, i.e. $\mathcal{H}_1, \mathcal{H}_2 \subset (\mathcal{H}_1 \oplus \mathcal{H}_2)$ and

$${}_1\langle\psi_1|\psi_2\rangle_2 = 0 \quad \forall |\psi_1\rangle_1 \in \mathcal{H}_1, |\psi_2\rangle_2 \in \mathcal{H}_2 \quad . \quad (43)$$

In order to simplify the notation the small indices indicating the Hilbert space a particular component belongs to, are left out.

Maybe the notion of a *direct sum* of Hilbert spaces appears rather unphysical to the reader, since everything in quantum mechanics is usually described in terms of direct products. But without explicitly using it, the direct sum is always present. For example, the fundamental space of quantum information theory, the space \mathbb{C}^2 of a single qbit is in fact the direct sum of two \mathbb{C} 's. The reason is that the *cartesian product* $\mathbb{C}^2 = \mathbb{C} \times \mathbb{C}$ can be embedded into the direct sum $\mathbb{C} \oplus \mathbb{C}$ by preserving the Hilbert space structure, since both components of \mathbb{C}^2 are mutually orthogonal. Hence the cartesian product and the direct sum are just different representations of the same principle: adding separate *levels* of an observable, i.e. combining properties of a system by a quantum mechanical *OR*. In many-letters theory the distinct degenerate levels of the length operator (see section IV B) are added, that is all.

The space $\mathcal{M}_{\mathcal{Q}}$ contains just *any* quantum message that can be composed from the quantum alphabet \mathcal{Q} by preparing each single letter state separately and then performing a unitary operation in the many-letter space on the entire message. Every Hilbert space $\mathcal{H}_{\mathcal{Q}}^n$ of block messages is a *subspace* of $\mathcal{M}_{\mathcal{Q}}$,

$$\forall n \in \mathbb{N} : \quad \mathcal{H}_{\mathcal{Q}}^n \subset \mathcal{M}_{\mathcal{Q}} \quad , \quad (44)$$

such that every Hilbert vector $|\psi\rangle \in \mathcal{H}_{\mathcal{Q}}^n$ is also an element of $\mathcal{M}_{\mathcal{Q}}$. That way, quantum information theory based on many-letter spaces *contains* quantum information theory based on block spaces, it can be viewed as a straight generalization of the latter. The many-letter space $\mathcal{M}_{\mathcal{Q}}$ is similar to the Fock space used in quantum optics or quantum statistics except that the states contained in $\mathcal{M}_{\mathcal{Q}}$ are neither symmetrized nor antisymmetrized. The *ordering* of the subspaces still matters (imagine a book written without ordering of the letters!). Therefore, the elements of $\mathcal{M}_{\mathcal{Q}}$ are neither Fermions nor Bosons, they are simply *quantum letters* and have to be realized by *distinguishable* quantum systems being separated in space or time or differing in some other observable property, such that their mutual overlap is neglectable.

Of course, the many-letter space can be restricted to a maximum number of letters,

$$\mathcal{M}_{\mathcal{Q}}^N := \bigoplus_{n=0}^N \mathcal{H}_{\mathcal{Q}}^n \quad , \quad (45)$$

due to a finite reservoir of available qbits. It is a subspace of the total many-letter space, $\mathcal{M}_{\mathcal{Q}}^N \subset \mathcal{M}_{\mathcal{Q}}$, and each many-letter message can be truncated to this subspace by the appropriate projector.

Quantum mechanics allows Alice to compose any superposition of block messages into a *general many-letter message*. Thus she uses the general message ensemble

$$|\Phi\rangle = \{[|\varphi\rangle, p(\varphi)] \mid |\varphi\rangle \in \Gamma\} \quad , \quad (46)$$

with the *source set* Γ of quantum messages being chosen with nonzero *a priori* probability $p(\varphi)$,

$$\Gamma = \{|\varphi\rangle \in \mathcal{M}_{\mathcal{Q}} \mid p(\varphi) > 0\} \quad . \quad (47)$$

Note that Γ may be an infinite set. The subspace spanned by the elements of Γ is the *source space* $\mathcal{M}_{\Gamma} \subset \mathcal{M}_{\mathcal{Q}}$,

$$\mathcal{M}_{\Gamma} := \text{Span}(\Gamma) \quad , \quad (48)$$

whose dimension $G := \dim \mathcal{M}_{\Gamma}$ may also be infinite. Equivalently, the message ensemble may be represented by a corresponding *message matrix* $\sigma \in \mathcal{S}(\mathcal{M}_{\mathcal{Q}})$, given by

$$\sigma = \sum_{\varphi \in \Gamma} p(\varphi) |\varphi\rangle \langle \varphi| \quad . \quad (49)$$

It is often more convenient to use the spectral decomposition of σ instead, given by

$$\sigma = \sum_{i=1}^G q_i |e_i\rangle \langle e_i| \quad , \quad (50)$$

where the $|e_i\rangle$'s form an orthonormal basis \mathcal{B}_{Γ} of the source space \mathcal{M}_{Γ} .

B. Length operator

To any classical letter $\mathbf{x} \in \mathcal{A}^+$ there is a length function $L : \mathcal{A}^+ \rightarrow \mathbb{N}$ mapping each letter \mathbf{x} to its length $L(\mathbf{x})$. Since the length of a quantum message is also an observable property (Bob has to measure the number of letter systems being engaged), there is a self-adjoint *length operator* \hat{L} acting on the many-letter space $\mathcal{M}_{\mathcal{Q}}$ with a spectral decomposition of mutually orthogonal projectors Π_n on $\mathcal{M}_{\mathcal{Q}}$, such that

$$\hat{L} = \sum_{n=0}^{\infty} n \Pi_n \quad , \quad (51)$$

with

$$\Pi_n \Pi_m = \delta_{nm} \Pi_n, \quad \sum_{n=1}^{\infty} \Pi_n = \mathbb{1} \quad . \quad (52)$$

The eigenspaces of the length operator are the block message spaces \mathcal{H}_Q^n , which are subspaces of the many-letter space \mathcal{M}_Q . Hence the eigenvalues of \hat{L} are degenerate by $K^n := \dim \mathcal{H}_Q^n = (\dim \mathcal{H}_Q)^n$. The projector Π_n onto the subspace \mathcal{H}_Q^n can be decomposed into mutually orthogonal product messages $|a^n\rangle$ of length n composed from a basis alphabet $\mathcal{B}_Q = \{|a\rangle\}_a$, where $|\mathcal{B}_Q^n| = |\mathcal{B}_Q|^n = K^n$ and $|a^0\rangle := |\cdot\rangle \in \mathcal{H}_Q^0$. The set of product messages of length n composed from the basis alphabet \mathcal{B}_Q is denoted by $\mathcal{B}_Q^n := \{|a^n\rangle\}_{a^n}$, the basis for the one-dimensional empty message space by $\mathcal{B}_Q^0 = \{|\cdot\rangle\}$. So the projector Π_n may be decomposed as

$$\Pi_n = \sum_{a^n} |a^n\rangle\langle a^n| \quad , \quad (53)$$

where we understand the sum as being performed over all quantum strings $|a^n\rangle \in \mathcal{B}_Q^n$ here and in the following. Using the basis

$$\mathcal{B}_Q^+ := \bigcup_{n=0}^{\infty} \mathcal{B}_Q^n \quad , \quad (54)$$

one arrives at the unity decomposition

$$\sum_{n=0}^{\infty} \sum_{a^n} |a^n\rangle\langle a^n| = \mathbb{1} \quad , \quad (55)$$

where the length operator becomes diagonal. Now Alice chooses a *general many-letter message* $|\varphi\rangle \in \Gamma \subset \mathcal{M}_Q$, whose decomposition in the basis \mathcal{B}_Q^+ thus reads

$$|\varphi\rangle = \sum_{n=0}^{\infty} \sum_{a^n} \varphi(a^n) |a^n\rangle \quad , \quad (56)$$

with its wave components given by

$$\varphi(a^n) := \langle a^n | \varphi \rangle \quad . \quad (57)$$

She sends her message to Bob using a quantum channel that is protected against decoherence of the basis vectors $|a^n\rangle$. That way, superpositions of these vectors are preserved and Bob receives the same state that Alice prepared. Now he measures the length of the message, obtaining random results with the *expected length* given by

$$L(\varphi) = \langle \varphi | \hat{L} | \varphi \rangle = \sum_{n=0}^{\infty} \sum_{a^n} |\varphi(a^n)|^2 n \quad , \quad (58)$$

whereas the *ensemble length* of the message σ is given by

$$L(\sigma) = \langle \Phi | \hat{L} | \Phi \rangle = \sum_{\varphi \in \Gamma} p(\varphi) L(\varphi) \quad (59)$$

$$= \sum_{\varphi \in \Gamma} \sum_{n=0}^{\infty} \sum_{a^n} p(\varphi) |\varphi(a^n)|^2 n \quad (60)$$

$$= \text{Tr}\{\sigma \hat{L}\} \quad . \quad (61)$$

As a generalization, we can define the *expected length* of any (pure or mixed) message, represented by a density matrix $\rho \in \mathcal{S}(\mathcal{M}_Q)$, by

$$L(\rho) := \text{Tr}\{\rho \hat{L}\} \quad . \quad (62)$$

Needless to say, the measurement of the length of a message will result in losing all quantum correlations between wave components of distinct length.

C. Random block messages

Alice now chooses block messages $|\varphi\rangle$ from any one of the subspaces $\mathcal{H}_Q^n \subset \mathcal{M}_Q$ with *a priori* probabilities $p(\varphi)$, i.e. she draws her messages from the ensemble

$$|\Phi\rangle = \{[|\varphi\rangle, p(\varphi)] \mid |\varphi\rangle \in \Gamma\} \quad , \quad (63)$$

where Γ is the set of block messages chosen with nonzero probability:

$$\Gamma := \{|\varphi\rangle \in \mathcal{H}_Q^n \mid p(\varphi) > 0, n = 0, 1, 2, \dots\} \quad , \quad (64)$$

The corresponding message matrix reads

$$\sigma = \sum_{\varphi \in \Gamma} p(\varphi) |\varphi\rangle\langle \varphi| \quad . \quad (65)$$

Every message $|\varphi\rangle$ drawn from the ensemble has a well-defined length $L(\varphi)$ because it is in one of the eigenspaces of the length operator, i.e. $\hat{L}|\varphi\rangle = L(\varphi)|\varphi\rangle$. Thus the message matrix of random block messages can be *block-diagonalized* into the convex combination of *block matrices* σ_n ,

$$\sigma = \sum_{n=0}^{\infty} \lambda_n \sigma_n \quad , \quad (66)$$

with the *length probabilities* λ_n , given by

$$\lambda_n := \sum_{L(\varphi)=n} p(\varphi) \quad , \quad (67)$$

such that

$$\sum_{n=0}^{\infty} \lambda_n = \sum_{n=0}^{\infty} \sum_{L(\varphi)=n} p(\varphi) = \sum_{\varphi \in \Gamma} p(\varphi) = 1. \quad (68)$$

Every block matrix has a definite length $\hat{L}\sigma_n = n\sigma_n$, hence it commutes with the length operator. So the average length of the ensemble reads

$$L(\sigma) = \sum_{n=0}^{\infty} \lambda_n n \quad . \quad (69)$$

We chose basis sets B_n of mutually orthogonal block messages $|e_{i_n}^n\rangle$ of length n , so that the block matrices become diagonal:

$$\sigma_n = \sum_{L(\varphi)=n} \sum_{i_n=1}^{K_n} |\varphi_{i_n}^n|^2 |e_{i_n}^n\rangle \langle e_{i_n}^n|, \quad (70)$$

with the wave components $\varphi_{i_n}^n := \langle e_{i_n}^n | \varphi \rangle$. Note that the block messages $|e_{i_n}^n\rangle$ are generally no product messages.

To Bob it appears as if Alice would send him states $|e_{i_n}^n\rangle$ of well-defined length n with the probability

$$P(e_{i_n}^n) = \lambda_n \sum_{L(\varphi)=n} |\varphi_{i_n}^n|^2. \quad (71)$$

A major advantage of using random block messages is that the length may be measured without disturbing the message.

D. Grand canonical messages

Grand canonical messages (or *random canonical messages*) are canonical messages of variable length (just like in thermodynamics, where grand canonical ensembles are canonical ensembles with variable particle number). Each classical letter $x^n \in \mathcal{A}^+$ of variable length n , composed from a classical alphabet \mathcal{A} is mapped to a product vector $|x^n\rangle \in \mathcal{H}_{\mathcal{Q}}^n$ and chosen by Alice with *a priori* probability $p(x^n)$. Alice thus draws her quantum messages from the ensemble

$$|\mathbf{X}\rangle = \{|x^n\rangle, p(x^n) \mid |x^n\rangle \in \Gamma, n = 0, 1, 2, \dots\}, \quad (72)$$

where the source set $\Gamma = \{|x^n\rangle \in \mathcal{H}_{\mathcal{Q}}^n \mid p(x^n) > 0, n = 0, 1, 2, \dots\}$, consists of canonical messages $|x^n\rangle$ of variable length $\hat{L}|x^n\rangle = n|x^n\rangle$, distributed by

$$p(x^n) := \lambda_n p(x_1) \cdots p(x_n), \quad (73)$$

where

$$\sum_x p(x) = 1, \quad \sum_{n=0}^{\infty} \lambda_n = 1. \quad (74)$$

The grand canonical message matrix has the form

$$\sigma = \sum_{n=0}^{\infty} \lambda_n \rho^{\otimes n}, \quad (75)$$

with the *block matrices*

$$\rho^{\otimes n} = \rho \otimes \cdots \otimes \rho, \quad (76)$$

and the *letter matrices*

$$\rho = \sum_x p(x) |x\rangle \langle x|. \quad (77)$$

Each block matrix has a definite length $\hat{L} \rho^{\otimes n} = n \rho^{\otimes n}$, so the *average length* of a grand canonical message ensemble is given by

$$L(\sigma) = \sum_{n=0}^{\infty} \lambda_n n. \quad (78)$$

Grand canonical messages can be viewed as a generalization of canonical messages, in that the length of a message is allowed to vary. Just as for every random block message, grand canonical messages are not disturbed by measuring the length operator.

We chose the basis sets $\mathcal{B}_{\mathcal{Q}} = \{|a\rangle\}$ so that the letter matrices become diagonal. The message matrix now reads

$$\sigma = \sum_{n=0}^{\infty} \sum_{x^n, a^n} \lambda_n p(x_1) \cdots p(x_n) |x^n(a^n)\rangle \langle a^n|, \quad (79)$$

with the wave components $x^n(a^n) := \langle a^n | x^n \rangle = \langle a_1 | x_1 \rangle \cdots \langle a_n | x_n \rangle$. To Bob it appears as if Alice would send him canonical messages $|a^n\rangle$ over the basis alphabet and of length n , composed from the basis alphabet $\mathcal{B}_{\mathcal{Q}}^n$ with the probability

$$q(a^n) = \langle a^n | \sigma | a^n \rangle \quad (80)$$

V. SUMMARY AND OUTLOOK

A framework has been worked out that makes the theoretical description of *many-letter states* possible, i.e. states consisting of arbitrary superpositions of quantum messages of distinct length. The space spanned by these states is the *many-letter space*, which is an infinite direct sum over all *block spaces*, i.e. finite dimensional Hilbert spaces containing quantum messages of fixed length. In the many-letter space a *length operator* is definable whose eigenspaces are the block spaces and where each eigenvalue is the number of letter systems forming the corresponding eigenspace.

The concept of many-letter messages can be applied to many topics of quantum information theory. Imagine a source of photons being sent sequentially, but whose number is controlled by the state of a quantum mechanical system. A superposition of input states will result in a superposition of distinguishable photon states of varying number forming a many-letter message whose length is a quantum mechanical observable with distinct values in superposition. Quantum communication, extended to the framework of many-letters, obtains new features. Quantum cryptography might also be affected (imagine an eavesdropper who is not allowed to measure the length of a message without disturbing it), as well as quantum computation (the output of a quantum algorithm can be regarded as a many-letter message). It is also interesting to study the entanglement of many-letter messages. Altogether, I hope that the presented concept will be helpful in many fields of quantum information theory.

VI. ACKNOWLEDGEMENTS

I would like to thank Jens Eisert, Timo Felbinger, Alexander Albus, and Shash Virmani for fruitful and intensive discussions about the topic of this paper.

- [1] D.J.C. MacKay.
Information theory, inference, and learning algorithms.
<http://wol.ra.phy.cam.ac.uk/mackay/itprnn/book.html>,
1995-2000.
- [2] J. Preskill. Lecture notes.
<http://www.theory.caltech.edu/people/preskill/ph219/>,
1997-1999.